



52 GTGCGATAGAGATTGACAGATATAGATGCTCAAGTATGAT 97  
|||||

seq\_name: gb\_pl2:AF161883

seq\_documentation\_block:

LOCUS AF161883 2182 bp mRNA PLN 10-SEP-1999  
DEFINITION Macadamia integrifolia vicillin precursor (AMP2) mRNA, AMP2-1  
allele, complete cds.  
ACCESSION AF161883  
VERSION AF161883.1 GI:5852871  
KEYWORDS  
SOURCE  
ORGANISM  
Macadamia integrifolia.  
Macadamia integrifolia.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
Proteaceae; Macadamia.

REFERENCE  
AUTHORS  
TITLE  
1 (bases 1 to 2182)  
Marcus, J.P., Goulter, K.C., Green, J.L. and Manners, J.M.  
A family of antimicrobial peptides is produced by processing of a  
7S globulin protein in Macadamia integrifolia kernels

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
2 (bases 1 to 2182)  
The Plant Journal (1999) In press

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
Submitted (24-JUN-1999) Cooperative Research Center for Tropical  
Plant Pathology, The University of Queensland, John Hines Building,  
QLD 4072, Australia

FEATURES  
source  
Location/Qualifiers

1..2182  
/organism="Macadamia integrifolia"  
/db\_xref="taxon:60698"  
/tissue\_type="nut kernel"  
1..2182  
/gene="AMP2"  
/allele="1"  
12..2012  
/gene="AMP2"  
/note="MAMP2 precursor; antimicrobial peptide family  
protein"  
/codon\_start=1  
/product="vicillin precursor"  
/protein\_id="AAD54244.1"  
/db\_xref="GI:5852872"  
/translation="MAINTSNLSCLFLSLFLSTVLSAESEFROEYBECKROCM  
QLETSQMRRCVQCKRREEDIDMSKYDNDPQDCCQCCQRCRQSGPQOQYC  
ORCKEICEEEDENRQDPOQOYECQCRQRETEPRMOTCCQRCERREYERK  
QQRKYEQOQDEDEKYEERKEEDNKRDPQREYDCRRCEQOEPQOQRCRE  
QQRQHRGQDMANPQRGSGRYEGEESQSNPYTFPERSLSTRFEBGHISYLEN  
YGRSKLRALKNRYLVLEANNPAFLPTHLADAILLVGGGALAKMHRDRESYN  
LECGDVRIPAGTFEYLNDNRNRLHIAKLTQISTPGQKEFFPAGQOPEYLSL  
FSKEILPAINTQTEKLRVGGQREGVITIASQOJRELTRDSESRHMLRGGES  
SRGPYVNSGENEPLYSNKYGAEVKPEYDROLQMDLSVETANTQSGSMGPEPNTS  
TKVYVVASGENEVLLEFAGINQNNHNEFLAGREKRVNLOQIEPQAMELAFAPRKEY  
GHVYVVSQGENEVLLEFAGINQNNHNEFLAGREKRVNLOQIEPQAMELAFAPRKEY  
EELFNSODESIFFPGRQHQOQSSRSTKQOQPLVSLDVGVF"

BASE COUNT 691 a 508 c 571 g 412 t  
ORIGIN

alignment\_scores:  
Quality: 180.00 Length: 32  
Ratio: 5.625 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-331-631-5\_COPY\_1\_32 x AF161883 ..

Align seg 1/1 to: AF161883 from: 1 to: 2182

1 GlnCysMetGlnLeuGluThrSerGlyGlnMetArgCysValSerGI 17  
|||||  
135 CATTGATGATGAGTGGAGACATCGACGATCGCTGGGTGTGAGTCA 184

17 nCysAspLysArgpHegluGluAspIleAspTrpSerIlystyAsp 32  
|||||  
185 GTGCGATAGAGATTGACAGATATAGATGCTCAAGTATGAT 230

seq\_name: gb\_pl2:AF161884

seq\_documentation\_block:

LOCUS AF161884 2182 bp mRNA PLN 10-SEP-1999  
DEFINITION Macadamia integrifolia vicillin precursor (AMP2) mRNA, AMP2-2  
allele, complete cds.  
ACCESSION AF161884  
VERSION AF161884.1 GI:5852873  
KEYWORDS  
SOURCE  
ORGANISM  
Macadamia integrifolia.  
Macadamia integrifolia.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;  
Proteaceae; Macadamia.

REFERENCE  
AUTHORS  
TITLE  
1 (bases 1 to 2182)  
Marcus, J.P., Goulter, K.C., Green, J.L. and Manners, J.M.  
A family of antimicrobial peptides is produced by processing of a  
7S globulin protein in Macadamia integrifolia

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
2 (bases 1 to 2182)  
The Plant Journal (1999) In press

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
Submitted (24-JUN-1999) Cooperative Research Center for Tropical  
Plant Pathology, The University of Queensland, John Hines Building,  
QLD 4072, Australia

FEATURES  
source  
Location/Qualifiers

1..2182  
/organism="Macadamia integrifolia"  
/db\_xref="taxon:60698"  
/tissue\_type="nut kernel"  
1..2182  
/gene="AMP2"  
/allele="2"  
12..2012  
/gene="AMP2"  
/note="MAMP2 precursor; antimicrobial peptide family  
member"  
/codon\_start=1  
/product="vicillin precursor"  
/protein\_id="AAD54245.1"  
/db\_xref="GI:5852874"  
/translation="MAINTSNLSCLFLSLFLSTVLSAESEFROEYBECKROCM  
QLETSQMRRCVQCKRREEDIDMSKYDNDPQDCCQCCQRCRQSGPQOQYC  
ORCKEICEEEDENRQDPOQOYECQCRQRETEPRMOTCCQRCERREYERK  
QQRKYEQOQDEDEKYEERKEEDNKRDPQREYDCRRCEQOEPQOQRCRE  
QQRQHRGQDMANPQRGSGRYEGEESQSNPYTFPERSLSTRFEBGHISYLEN  
YGRSKLRALKNRYLVLEANNPAFLPTHLADAILLVGGGALAKMHRDRESYN  
LECGDVRIPAGTFEYLNDNRNRLHIAKLTQISTPGQKEFFPAGQOPEYLSL  
FSKEILPAINTQTEKLRVGGQREGVITIASQOJRELTRDSESRHMLRGGES  
SRGPYVNSGENEPLYSNKYGAEVKPEYDROLQMDLSVETANTQSGSMGPEPNTS  
TKVYVVASGENEVLLEFAGINQNNHNEFLAGREKRVNLOQIEPQAMELAFAPRKEY  
GHVYVVSQGENEVLLEFAGINQNNHNEFLAGREKRVNLOQIEPQAMELAFAPRKEY  
EELFNSODESIFFPGRQHQOQSSRSTKQOQPLVSLDVGVF"

BASE COUNT 684 a 510 c 572 g 416 t  
ORIGIN

alignment\_scores:  
Quality: 180.00 Length: 32  
Ratio: 5.625 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-331-631-5\_COPY\_1\_32 x AF161884 ..

Align seg 1/1 to: AF161884 from: 1 to: 2182

1 GlnCysMetGlnLeuGluThrSerGlyGlnMetArgCysValSerGI 17  
|||||

```

135 CAATGCATGACGTGGAGACATCAGCCAGCATCGTCGGTGTGAGTGA 184
|||||
17 nCySaapLySArgPhgUGLuAspIleAspTrpSerLysTYrASP 32
|||||
185 GTCCGATAAGAGATTGAAGAGCATATGATGCTTAAGTAGAT 230
|||||

seq_name: gb_om:BTRNAT2

seq_documentation_block:
LOCUS BTRNAT2 2053 bp mRNA MAM 07-JUN-1995
DEFINITION B.taurus mRNA for thrombospondin (partial) 2055 bp.
ACCESSION X87619
VERSION X87619.1 GI:860884
KEYWORDS thrombospondin; TSP-1 gene.
SOURCE Bos taurus.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Artiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE
AUTHORS Zafar,R.S., Moll,Y.D., Womack,J.F. and Walz,D.A.
TITLE Cloning and sequencing of bovine thrombospondin stimulatory effect
of TGF-beta
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2053)
AUTHORS Zafar,R.S.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-1995) R.S. Zafar, Wayne State University, Dept of
Physiology, 540 E. Canfield, Detroit MI 48202, CANADA
FEATURES
source
1..2053
/organism="Bos taurus"
/db_xref="taxon:9913"
/tissue_type="aortic endothelium"
/clone_lib="lambda gcl0"
/clone_pdb="pcem-32f(-)"
1..6
/note="EcoI cloning site"
misc_feature
180..233
/gene="TSP-1"
180..233
/gene="TSP-1"
180..>233
/gene="TSP-1"
/function="adhesive glycoprotein"
/codon_start=1
/product="thrombospondin"
/protein_id="CAA60951.1"
/db_xref="GI:860885"
/db_xref="SPREMBL:Q28179"
/translation="MGLAWMGVILLHACGS"
BASE COUNT 453 a 625 c 593 g 382 t
ORIGIN

alignment_scores:
Quality: 64.50 Length: 26
Ratio: 3.395 Gaps: 1
Percent Similarity: 73.077 Percent Identity: 46.154

alignment_block:
US-09-331-631-5_COPY_1_32.x BTRNAT2 ..
Align seg 1/1 to: BTRNAT2 from: 1 to: 2053

7 ThrSerGlyGlnMetArgArgCys...ValSerGlnCysAspLysArgPhe 22
:::|||| ||| ||| ::::::::::::::::::::|||
1424 TCCTCTGTCAACAAGCGGACTGCCACATCCAGAGAGTGACAAAGAGATT 1473
|||||
22 eGluGluAspIleAspTrpSerLysTyR 31
|::::::::::| |:::::
1474 TAACAAGAGATGGCGCTGGAGCCACATGG 1501

```

[illegible]

seq\_name: gb\_pr2:HSTHRMR

seq\_documentation\_block:

LOCUS HSTHRMR 4434 bp mRNA

PRI

21-MAR-1995

DEFINITION Human mRNA for thrombospondin.

ACCESSION X04665.1

GI:37137

KEYWORDS glycoprotein; signal peptide; thrombospondin.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 4434)

Lawler, J. and Hynes, R.O.

The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologues with several different proteins

J. Cell Biol. 103 (5), 1635-1648 (1986)

87057617

JOURNAL

MEDLINE

COMMENT

Three types of repeating amino acid sequence are present in thrombospondin. The first is 57 amino acids long and shows homology with circumsporozoite protein from Plasmodium falciparum. The second is 50-60 amino acids long and shows homology with epidermal growth factor precursor. The third occurs as a continuous eightfold repeat of a 38-residue sequence; structural homology with parvalbumin and calmodulin indicates that these repeats constitute the multiple calcium-binding sites of thrombospondin. Data kindly reviewed (15-SEP-1987) by Lawler J.

FEATURES

SOURCE

1. 4434

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/cell\_type="umbilical vein endothelial cells"

76. 3588

/note="precursor polypeptide (AA -18 to 1152)"

/codon\_start=1

/protein\_id="CAA28370.1"

/db\_xref="GI:37138"

/db\_xref="SMISS-PROT:P07996"

/translation="MGLAMGLVLLHAGCGSNRIPESGSDNSVPDIFELTGAARRKSGRLVKGDPSSPAFRIDANLIPVPDPRKFDLVDAVTRTEKGLILLASLQMKRTKSTLALERKDHSGVFSVSNKAGLTLSDLVQKQHYVSEALNATGQKSTLLEQEDRAQLIDICEKEMNELDVPVQSVFTRDASLARLHAGVNDNFQALQVREVFETPEDILRNKCGSSSTVLLTDNNVYNSSPAIRNTYGHKTKDLQALCGSCDELSMWLELRGLRTVTTLODSIRKYTEENKLANELRRPILCYNGOYRNNEMTVDSCTECHQNSVTICKKVCSPIMPCSNATVAGDECCPCPMPSDADDMGSPSWSCSTGTCNGITQRRSCDSLNRCEGSSVQRTCHTQEDCKRFKODGGRSHKSPWSCSTGCGDVITRILNCSFSPOMNGKPCGEARETKACKKDCPLNGMGWSPSPDICSVC

GGGVQKRSRLCNPTPOFGKDCVQDVENQICKKDCPLNGMGWSPSPDICSVC

DGSGWKGACPPGSGNGIOCTYDECKEVPACNHNKNCNTLGYSDPMRCEKPCYTGSOPFGGVEHATANKOVCKRPNPCTGTGTDCKNKKACNTLGYSDPMRCEKPCYAGNGLIGEDTDLGMPNENLVYANATYHCKKNCPLNPGSGEDDQKQISGACDDDDNDKIPDDNDNCPFHNPQOYVDRDQVDCRDNCPYNNPQATDNNGEEDACA

ADIDGGLNERNKQYVYNDQDVTDDGVDGDCDNCPLSHNDQDLSDDSDGTC

DNNDIDEDGHQNNLDNCPYNNPQANQDHDGDCADHDNDGIPDDKNCRLV

NPDKDSGDGRGACKDEFDHSDPIDIDCPENVIDISERFERFQMIPLDPSGTQ

NDPMVYVHOKELVQYVNCDPGLAVGDEFNADVSGTFPIINERDDYAGVEYQ

SSRFYVVMKQVQSYWDTNPTPAQCGSGLSVYVSTGPGHLNRLNHLNHTNIG

OVRTLMHDPRTIGKDFATYRWRLSHRPKTGFIVYMEGKTIADSGPIITKATG

RLGLFVSQEWVFSDLYECRDP

76. 129

/product="put. signal peptide (AA -18 to -1)"

130. 3585

/product="mature peptide (AA 1-1152)"

817. 825

/note="pot. N-glycosylation site"

1153. 1161

/note="pot. N-glycosylation site"

1636. 1644

/note="pot. N-glycosylation site"

2197. 2205

/note="pot. N-glycosylation site"

misc\_feature 3226..3234

/note="pot. N-glycosylation site"

misc\_feature 3274..3282

/note="pot. N-glycosylation site"

BASE COUNT 1139 a 1185 c 1177 g 933 t

ORIGIN

alignment\_scores:

Quality: 64.50 Length: 26

Ratio: 3.395 Gaps: 1

Percent Similarity: 73.077 Percent Identity: 46.154

alignment\_block:

US-09-331-631-5\_COPY\_1\_32 x HSTHRMR

Align seg 1/1 to: HSTHRMR from: 1 to: 4434

7 ThrSerGlyGlnMetArgArgGys...ValSerGlnCysAspLysArgph 22

1321 TCCTGGTCCAGACAGCGACTGCACATTCAGAGTGTGACACAGAGATT 1370

22 eGLuGLuAspLLeAspTrpSerLysTyr 31

1371 TAAACGAGTGTGCTGAGCCACTGG 1398

seq\_name: gb\_cm:AB005287

seq\_documentation\_block:

LOCUS AB005287 5269 bp mRNA

DEFINITION Bos taurus mRNA for thrombospondin 1, complete cds.

ACCESSION AB005287

VERSION AB005287.2 GI:4630795

KEYWORDS thrombospondin 1.

SOURCE Bos taurus

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae;

Bovinae; Bos.

REFERENCE 1 (sites)

Ueno, A., Yamashita, K., Nagata, T., Tsurumi, C., Miwa, Y., Kitamura, S.

and Inoue, H.

cDNA cloning of bovine thrombospondin 1 and its expression in

odontoblasts and predentin

Biochim. Biophys. Acta 1382 (1), 17-22 (1998)

98173773

2 (bases 1 to 5269)

Ueno, A.

Direct Submission

Submitted (30-JUN-1997) to the DDBJ/EMBL/Genbank databases.

Atemichi Ueno, The University of Tokushima, Department of

Biochemistry, School of Dentistry, Kuramoto-cho 3-18-15, Tokushima,

Tokushima 770, Japan (E-mail:atemichi@dent.tokushima-u.ac.jp,

Tel:+81-886-33-7326, Fax:+81-886-31-4215)

On Apr. 21, 1999 this sequence version replaced gi:2244706.

Location/Qualifiers

1. 5269

/organism="Bos taurus"

/strain="Holstein"

/db\_xref="taxon:9913"

/cell\_type="odontoblast"

/clone="PBTS1-28"

/dev\_stage="two weeks after birth"

/sex="male"

/tissue\_type="anterior tooth"

148. 3660

/codon\_start=1

/product="thrombospondin 1"

/protein\_id="BA21115.1"

/db\_xref="GI:2244707"

/translation="MGLAMGLVLLHAGCGSNRIPESGSDNSVPDIFELTGAARRKSGRLVKGDPSSPAFRIDANLIPVPDPRKFDLVDAVTRTEKGLILLASLQMKRTK

STLALERKDHSGVFSVSNKAGLTLSDLVQKQHYVSEALNATGQKSTLLEQEDRAQLIDICEKEMNELDVPVQSVFTRDASLARLHAGVNDNFQALQVREVFETPEDILRNKCGSSSTVLLTDNNVYNSSPAIRNTYGHKTKDLQALCGSCDELSMWLELRGLRTVTTLODSIRKYTEENKLANELRRPILCYNGOYRNNEMTVDSCTECHQNSVTICKKVCSPIMPCSNATVAGDECCPCPMPSDADDMGSPSWSCSTGTCNGITQRRSCDSLNRCEGSSVQRTCHTQEDCKRFKODGGRSHKSPWSCSTGCGDVITRILNCSFSPOMNGKPCGEARETKACKKDCPLNGMGWSPSPDICSVC

GGGVQKRSRLCNPTPOFGKDCVQDVENQICKKDCPLNGMGWSPSPDICSVC

DGSGWKGACPPGSGNGIOCTYDECKEVPACNHNKNCNTLGYSDPMRCEKPCYTGSOPFGGVEHATANKOVCKRPNPCTGTGTDCKNKKACNTLGYSDPMRCEKPCYAGNGLIGEDTDLGMPNENLVYANATYHCKKNCPLNPGSGEDDQKQISGACDDDDNDKIPDDNDNCPFHNPQOYVDRDQVDCRDNCPYNNPQATDNNGEEDACA

ADIDGGLNERNKQYVYNDQDVTDDGVDGDCDNCPLSHNDQDLSDDSDGTC



```

Align seg 1/1 to: HSTS from: 1 to: 5722
7 ThisSerGlyGlnMetArgCys...ValSerGlnCysAspLysArgPh 22
:::||||| ||| ||| ||| :::::::::::||||| ||| ||| |||
1357 TCCTCGTTCACGACGACGACCTGCGACATTCAGAGGTGACAAAGATT 1406
22 eGIUGLUAspIleAspTrpSerLysTyr 31
:::||||| ||| ||| ||| ::
1407 TAAACAGATGCTGCTGAGCAGCATTG 1434

seq_name: gb_ov:XELSTAGE

seq_documentation_block:
LOCUS XELSTAGE 3522 bp mRNA VRT 12-JUN-1993
DEFINITION Xenopus laevis mRNA sequence.
ACCESSION L04278
VERSION L04278.1 GI:295541
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis stage 45 cDNA to mRNA.
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodidae;
Xenopus.
REFERENCE
1 (bases 1 to 3522)
Urry,L.A., Ramos,J., Duquette,M., Desimone,D.W. and Lawler,J.
Cloning, characterization and expression of thrombospondin-1 in
Xenopus laevis embryos
JOURNAL
Unpublished (1992)
FEATURES
Location/Qualifiers
1..3522
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/dev_stage="stage 45"
BASE COUNT 1031 a 786 c 884 g 821 t
ORIGIN

alignment_scores:
Quality: 62.50 Length: 26
Ratio: 3.289 Gaps: 1
Percent Similarity: 73.077 Percent Identity: 46.154

alignment_block:
US-09-331-631-5_COPY_1_32 x XELSTAGE ..
Align seg 1/1 to: XELSTAGE from: 1 to: 3522
7 ThisSerGlyGlnMetArgCys...ValSerGlnCysAspLysArgPh 22
:::||||| ||| ||| ||| :::::::::::||||| ||| ||| |||
1255 TCCTCGTTCACGACGACGACCTGCGACATTCAGAGGTGACAAAGATT 1304
22 eGIUGLUAspIleAspTrpSerLysTyr 31
:::||||| ||| ||| ||| ::
1305 TAAACAGACGCTGCTGAGCAGCATTG 1332

seq_name: gb_hcg14:AC020788

seq_documentation_block:
LOCUS AC020788 197071 bp DNA HTG 19-FEB-2000
DEFINITION Homo sapiens chromosome 3 clone RP11-162J10. *** SEQUENCING IN
PROGRESS ***, 113 unordered pieces.
ACCESSION AC020788
VERSION AC020788.2 GI:7007946
KEYWORDS HTG: HTGS_PHASE1.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 197071)
Muzny,D.M., Adams,C., Bailey,M., Barbara,J., Blankenburg,K.,
Bodoty,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
Forcun-Tansey,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,
Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,
Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,
Kelly,S., Kondewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,
Lichtarge,O., Liu,J., Liu,W., Logan,O., Lozado,R.J., Lu,J.,
Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mel,G., Morgan,M.,
Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S.,
Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,
Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,
Shah,E., Shen,H., Simon,M., Sparks,A., Stamps,A., Sugang,R.,
Taber,P., Taylor,T., Vasquez,L., Vinson,R., Vo,Q., Wahbah,M.,
Watlington,S., Weinstein,G., Weinstein,I.R., Williamson,A.,
Worley,K., Wren,J., Wrenford,G., Yu,W., Zhou,X., Nelson,D. and
Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 197071)
Worley,K.C.
Direct Submission
Submitted (10-JUN-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Feb 19, 2000 this sequence version replaced gi:5686414.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HMXE
Center clone name: RP11-162J10
----- Summary Statistics
Sequencing vector: M13; L08821
Chemistry: Dye-terminator Big Dye; version 0.980611
Assembly program: Phrap; version 0.980611
Consensus quality: 103326 bases at least Q40
Consensus quality: 135876 bases at least Q30
Consensus quality: 151983 bases at least Q20
Estimated insert size: 165698; sum-of-contigs estimation
Quality coverage: 1.2x in Q20 bases; sum-of-contigs estimation

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 113 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 785: contig of 785 bp in length
* 786 805: gap of unknown length
* 1571 1571: contig of 766 bp in length
* 1572 1591: gap of unknown length
* 1592 2373: contig of 782 bp in length
* 2374 2393: gap of unknown length
* 2394 3161: contig of 768 bp in length
* 3162 3181: gap of unknown length
* 3182 3945: contig of 764 bp in length
* 3946 3965: gap of unknown length
* 3966 4759: contig of 794 bp in length
* 4760 4779: gap of unknown length
* 4780 5649: contig of 870 bp in length
* 5650 5669: gap of unknown length
* 5670 6452: contig of 783 bp in length
* 6453 6472: gap of unknown length
* 6473 7253: contig of 781 bp in length
* 7254 7273: gap of unknown length
* 7274 8074: contig of 801 bp in length
* 8075 8094: gap of unknown length
* 8095 8863: contig of 769 bp in length
* 8864 8883: gap of unknown length
* 8884 9669: contig of 786 bp in length

```





TITLE Mueller, R.  
New lessons for combinatorial biosynthesis from myxobacteria. The  
myxobactiazol biosynthetic gene cluster of *Stigmatella aurantiaca*  
DM4/3-1

JOURNAL J. Biol. Chem. 274 (52), 37391-37399 (1999)

MEDLINE 20069734

REFERENCE 2 (bases 1 to 42603)

AUTHORS Silakowski, B., Schaefer, H.U., Ehret, H., Kunze, B., Weinig, S.,  
Nordstedt, G., Brandt, P., Bloeker, H., Hoeffle, G., Beyer, S. and  
Mueller, R.

TITLE Direct Submission

JOURNAL Submitted (20-SEP-1999) NBI/MX, GBF, Mascheroder Weg 1,  
Braunschweig 38124, Germany

FEATURES  
source  
1. .42603  
/organism="Stigmatella aurantiaca"  
/strain="DM4/3-1"  
/db\_xref="taxon:41"  
1. .834  
/gene="mtaa"  
1. .834  
/gene="mtaa"  
/note="Pant transferase"  
/codon\_start=1  
/transl\_table=11  
/product="Mtaa"  
/protein\_id="AA19809.1"  
/db\_xref="GI:6635394"  
1388. .13399  
/gene="mtab"  
/note="PKS"  
1388. .13399  
/gene="mtab"  
/note="PKS: Mtab loading domain and module 1: polyketide  
synthase domains: ACP, KS, AT, DH, ER, KR, ACP; Mtab  
module 2: polyketide synthase domains: KS, AT, DH, KR,  
ACP"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="AA19810.1"  
/db\_xref="GI:6635395"

gene  
/transl\_table=11  
/product="Mtab"  
/protein\_id="AA19810.1"  
/db\_xref="GI:6635396"  
13537. .17409  
/gene="mtac"  
/note="mtac"  
/note="peptide synthetase; peptide synthetase domains:  
Heterocyclisation, adenylation, PCP note: motif with weak  
homology to putative NAD binding proteins from *Pyrococcus*  
species located behind PCP; this motif may be involved  
in-thiazole formation from the proposed thiazole  
intermediate of the biosynthesis of myxobactiazol"  
/codon\_start=1  
/transl\_table=11  
/product="Mtab"  
/protein\_id="AA19811.1"  
/db\_xref="GI:6635397"

13537. .17409  
/gene="mtac"  
/note="mtac"  
/note="peptide synthetase; peptide synthetase domains:  
Heterocyclisation, adenylation, PCP note: motif with weak  
homology to putative NAD binding proteins from *Pyrococcus*  
species located behind PCP; this motif may be involved  
in-thiazole formation from the proposed thiazole  
intermediate of the biosynthesis of myxobactiazol"  
/codon\_start=1  
/transl\_table=11  
/product="Mtab"  
/protein\_id="AA19811.1"  
/db\_xref="GI:6635397"





```
* *
* 11065 11998: gap of unknown length
* *      contig of 934 bp in length
* *      gap of unknown length
* 11999 12918: contig of 920 bp in length
* *      gap of unknown length
* 12919 13842: contig of 924 bp in length
* *      gap of unknown length
* 13843 14750: contig of 908 bp in length
* *      gap of unknown length
* 14751 15703: contig of 953 bp in length
* *      gap of unknown length
* 15704 16635: contig of 932 bp in length
* *      gap of unknown length
* 16636 17572: contig of 937 bp in length
* *      gap of unknown length
* 17573 18510: contig of 938 bp in length
* *      gap of unknown length
* 18511 19442: contig of 932 bp in length
* *      gap of unknown length
* 19443 20365: contig of 923 bp in length
* *      gap of unknown length
* 20366 21284: contig of 919 bp in length
* *      gap of unknown length
* 21285 22179: contig of 895 bp in length
* *      gap of unknown length
* 22180 23144: contig of 965 bp in length
* *      gap of unknown length
* 23145 24060: contig of 916 bp in length
* *      gap of unknown length
* 24061 25030: contig of 970 bp in length
* *      gap of unknown length
* 25031 25949: contig of 919 bp in length
* *      gap of unknown length
* 25950 26914: contig of 965 bp in length
* *      gap of unknown length
* 26915 27855: contig of 941 bp in length
* *      gap of unknown length
* 27856 28764: contig of 909 bp in length
* *      gap of unknown length
* 28765 29706: contig of 942 bp in length
* *      gap of unknown length
* 29707 30657: contig of 951 bp in length
* *      gap of unknown length
* 30658 31593: contig of 936 bp in length
* *      gap of unknown length
* 31594 32536: contig of 943 bp in length
* *      gap of unknown length
* 32537 33459: contig of 923 bp in length
* *      gap of unknown length
* 33460 34397: contig of 938 bp in length
* *      gap of unknown length
* 34398 35299: contig of 902 bp in length
* *      gap of unknown length
* 35300 36229: contig of 930 bp in length
* *      gap of unknown length
* 36230 37172: contig of 943 bp in length
* *      gap of unknown length
* 37173 38122: contig of 950 bp in length
* *      gap of unknown length
* 38123 39048: contig of 926 bp in length
* *      gap of unknown length
* 39049 40013: contig of 965 bp in length
* *      gap of unknown length
* 40014 40927: contig of 914 bp in length
* *      gap of unknown length
* 40928 41851: contig of 924 bp in length
* *      gap of unknown length
* 41852 42766: contig of 915 bp in length
* *      gap of unknown length
* 42767 43716: contig of 950 bp in length
* *      gap of unknown length
* 43717 44632: contig of 916 bp in length
* *      gap of unknown length
```

```
* *      44633 45540: contig of 908 bp in length
* *      gap of unknown length
* *      45541 46457: contig of 917 bp in length
* *      gap of unknown length
* *      46458 47425: contig of 968 bp in length
* *      gap of unknown length
* *      47426 48365: contig of 940 bp in length
* *      gap of unknown length
* *      48366 49294: contig of 929 bp in length
* *      gap of unknown length
* *      49295 50219: contig of 925 bp in length
* *      gap of unknown length
* *      50220 51142: contig of 923 bp in length
* *      gap of unknown length
* *      51143 52066: contig of 924 bp in length
* *      gap of unknown length
* *      52067 53020: contig of 954 bp in length
* *      gap of unknown length
* *      53021 53920: contig of 900 bp in length
* *      gap of unknown length
* *      53921 54875: contig of 955 bp in length
* *      gap of unknown length
* *      54876 55769: contig of 894 bp in length
* *      gap of unknown length
* *      55770 56697: contig of 928 bp in length
* *      gap of unknown length
* *      56698 57628: contig of 931 bp in length
* *      gap of unknown length
* *      57629 58568: contig of 940 bp in length
* *      gap of unknown length
* *      58569 59470: contig of 902 bp in length
* *      gap of unknown length
* *      59471 60424: contig of 954 bp in length
* *      gap of unknown length
* *      60425 61383: contig of 959 bp in length
* *      gap of unknown length
* *      61384 62336: contig of 953 bp in length
* *      gap of unknown length
* *      62337 63258: contig of 922 bp in length
* *      gap of unknown length
* *      63259 64141: contig of 883 bp in length
* *      gap of unknown length
* *      64142 65045: contig of 904 bp in length
* *      gap of unknown length
* *      65046 66008: contig of 963 bp in length
```

```
alignment_scores:
  Quality: 61.50      Length: 32
  Ratio: 2.674      Gaps: 1
  Percent Similarity: 71.875      Percent Identity: 37.500
```

```
alignment_block:
US-09-331-631-5_copy_1_32 x AC019039/rev ..
```

```
Align seg 1/1 to reverse of: AC019039 from: 1 to: 162168
```

```
2 CysMetcInleuGlutHrSerGlyGlnMetArgGArgCysValserGlnCy 18
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
33263 TGCATATCTCTCTCCCTGCTGCGCAATGTAATAATGACTCCCAATC 33214
18 sAspLYsArgPhcGlnGluAspIleAspTrpSer...LysTyrAsp 32
::: :::::||||:||||:||||:||||:||||:||||:||||:||||:
33213 TGAATGCGCATGGCCACGTGACTTGTGGCCGCAATAATATAGAG 33168
```

```
seg_name: gp_pr4:AC011331
```

```
seg_documentation_block:
```

```
LOCUS AC011331 175278 bp DNA PRI 06-OCT-1999
DEFINITION Genomic Sequence for Homo sapiens clone H_NH0514011, chromosome 18,
complete sequence.
ACCESSION AC011331
VERSION AC011331.1 GI:6013504
KEYWORDS HTG.
```

```

SOURCE          human.
ORGANISM        Homo sapiens
REFERENCE
AUTHORS         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
                Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 175278)
Huang, E.N., Nascimento, L.U., de la Bastide, M., Vil, D.M.,
Preston, R., Matero, A., Shah, R., O'Shaughnessy, A.,
Shkher, M., Schütz, K., See, L.H., Swaby, I., Habermann, K.,
Dechla, N.N. and McCombie, R.W.
Genomic Sequence for Homo sapiens clone H_NH0514011, Chromosome 18,
complete sequence
JOURNAL
REFERENCE       Unpublished
AUTHORS         2 (bases 1 to 175278)
                Huang, E.N., Nascimento, L.U., de la Bastide, M., Vil, D.M.,
                Spiegel, L.A., Preston, R.R., Matero, A., Shah, R., O'Shaughnessy, A.,
                Rodríguez, M., Shekhar, M., Schütz, K., See, L.H., Swaby, I.,
                Habermann, K., Dechla, N.N. and McCombie, R.W.
                Direct Submission
JOURNAL
COMMENT         Submitted (06-OCT-1999) Uta Annenberg Hazen Genome Sequencing
                Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
                Harbor, NY 11724, USA
                At location 59145 there is a subclone variation in which one
                subclone calls a 'G' and all the rest call 'A'. From 94100-94174
                there is single-strand/single chemistry. At 148425 there is a
                subclone variation in which one subclone calls 'T' and all others
                call 'G'.
FEATURES
SOURCE          Location/Qualifiers
                1..175278
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="18"
                /clone="H_NH0514011"
BASE COUNT      46991 a 36801 c 37956 g 53530 t
ORIGIN
alignment_scores:
                Quality: 61.50      Length: 32
                Ratio: 2.674      Gaps: 1
                Percent Similarity: 71.875      Percent Identity: 37.500
alignment_block:
US-09-331-631-5-COPY_1-32 x AC011331/rev ..
Align seg 1/1 to reverse of: AC011331 from: 1 to: 175278
                2 CysMetGlnLeuGlnThrSerGlyGlnMetArgCysValSerGlnCy 18
                |||:::||||:::|||||:::|||||:::|||||:::|||||:::|||||
54662 TGCATATCTCTCTCTCTCTCTCTGCGCATGTGTAATAATGACTTCCCAATC 54613
                18 sasplysArgpheGlnGluAspIleAspTyrPseR...LysTyrAsp 32
                :::::||||:::|||||:::|||||:::|||||:::|||||:::|||||
54612 TGAATGGCATGGCCACGCGACTTGCTTTGGCGCAGTAATAATATGAG 54567
seq_name: gb_1n1:CELC48E7
seq_documentation_block:
LOCUS          CELC48E7 37701 bp DNA INV 23-APR-1997
DEFINITION    Caenorhabditis elegans cosmid C48E7.
ACCESSION     AF000262
VERSION       AF000262.1 GI:1947126
KEYWORDS
SOURCE        Caenorhabditis elegans strain-Bristol N2.
ORGANISM      Caenorhabditis elegans
REFERENCE     Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
AUTHORS       Rhabditina; Rhabditicoidea; Rhabditidae; Pelodierinae; Caenorhabditis.
1 (bases 1 to 37701)
Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M.,
Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J.,
Coulson, A., Craxton, M., Dear, S., Du, Z., Dunabin, R., Favello, A.,
Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M.,
Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laister, N.,
Latterle, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B.,

```

```

TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

O'Callaghan,M., Parsons,J., Percy,C., Riffen,L., Roopira,A.,
Saunders,D., Showkneen,R., Smalton,N., Smith,A., Sonhammer,E.,
Staden,R., Sulston,J., Thierry-Mieg,J., Thomas,K., Vaudin,M.,
Vaughan,R., Waterston,R., Watson,A., Weinstock,L.,
Wilkinson-Sproat,J. and Wohlman,P.
2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans
Nature 368 (6466), 32-38 (1994)
94150718
2 (bases 1 to 37701)
Mamsley,P. and Kramer,J.
The sequence of C. elegans cosmid C487
Unpublished (1997)
3 (bases 1 to 37701)
Waterston,R.
Direct Submission
Submitted (18-APR-1997)
Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RO, England
e-mail: rtw@nematoce.wustl.edu and jes@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone.
It may be shorter because we only sequence overlapping sections
once, or longer because we provide a small overlap between
neighboring submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate
chemistry; an attempt was made to resolve all sequencing problems,
such as compressions and repeats; all regions were covered by
sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The 5' cosmid is T08B2, 700 bp overlap; 3' cosmid is C10G11, 700 bp
overlap. Actual start of this cosmid is at base position 1 of
CEC48E7. Actual end is at 6598 of CEC10G11

NOTES:

Coding sequences below are predicted from computer analysis, using
the program GeneFinder (P. Green and L. Hillier, ms in preparation).
Location/Qualifiers
1. 37701
/organism="Caenorhabditis elegans"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="I"
/clone="C48E7"
121. 624
/clone="C48E7.7"
join(121..200,246..624)
/gene="C48E7.7"
/codon_start=1
/feature="not experimental"
/protein_id="FAB52932.1"
/db_xref="GI:1947127"
/translaton="MOSINILFAMLLIAPLYNGDFNAVATATKVTEDATVALIETT
AKAVALEPTEPEVQTEAVEETFEAVVEATQETKDAETAPVDVPPANNIEATTEG
ASPPSSSTVAPAVTSSVDDLGETSTINAKAVNMGTFEVPVMDVALIQ"
4876..14780
/clone="C48E7.6"
join(4876..4956,5006..5146,5384..5730,5773..5907,
6158..6368,6463..6861,6835..7012,7062..7379,7606..7829,
7925..8052,8102..8334,8566..8755,8802..8980,9025..9205,
9256..9411,9456..9535,9715..9783,9857..9991,10484..10611)

```



AUTHORS Lin, X., Kaul, S., Rounsley, S.D., Shea, T.P., Benito, M.-I., Town, C.D., Fujii, C.Y., Mason, T.M., Bowman, C.L., Barnstead, M.E., Feldhym, T.V., Buell, C.R., Ketchum, K.A., Lee, J.J., Rinning, C.M., Koo, H., Moffat, K.S., Cronin, L.A., Shen, M., Vanaken, S.E., Umayam, L., Tallon, L.J., Gill, J.E., Adams, M.D., Carrera, A.J., Creasy, T.H., Goodman, H.M., Somerville, C.R., Copenhagen, G.P., Preuss, D., Nierman, W.C., White, O., Eisen, J.A., Salzberg, S.L., Fraser, C.M. and Venter, J.C.

TITLE Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*

JOURNAL Nature 402 (6763), 761-768 (1999)

MEDLINE 20083487

REFERENCE 2 (bases 1 to 84254)

AUTHORS Lin, X.

JOURNAL Direct Submission

COMMENT Submitted (13-DEC-1999) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA

On Dec 17, 1999 this sequence version replaced gi:2702261. The sequence and annotation of chromosome 2 were merged from those of the individual clones on this chromosome after removing overlaps. For detailed information, please see the TIGR web site (<http://www.tigr.org/db/at/at.html>).

Genes were identified by a combination of three methods: Gene prediction programs including GRAIL, GeneFinder (Phil Green, (<http://Arthur.epm.ornl.gov/pub/xgrail/>), GeneScan (Chris Burge, <http://genomic.stanford.edu/GENSCAN.html>), and NetPlantGene (<http://www.cbs.dtu.dk/services/NetGene2/>), searches of the complete sequence against a peptide database and plant EST databases at TIGR, and manual curations based on those analyses. Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by two or more gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Genes are numbered from the top to the bottom of the chromosome.

We thank the CSHL/MashU/ABI consortium for sequencing BAC clones F6P23, F5J6, T17A5, and T13L16, the ESSA group for sequencing clone F13D4, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards and Satoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Lixin Zhou, Hanif Khalak, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: [atltigr.org](mailto:atltigr.org).

FEATURES

source 1.84254 Location/Qualifiers

organism="Arabidopsis thaliana"

cultivar="Columbia"

db\_xref="taxon:3702"

chromosome="2"

complement(1..1230)

note="Sequence from clone F24L7"

misc\_feature 297..343

repeat\_region /rpt\_family="(GAA)n"

mrna <331..>1332

gene /gene="At2g32820"

<331..>1332

/gene="At2g32820"

/note="F24L7.4: predicted by genefinder and genscan"

CDS 331..1332

/gene="At2g32820"

/note="hypothetical protein"

/codon\_start=1

/protein\_id="BAB91965.2"

/db\_xref="GI:6598386"

/translation="WMKIAIATEETDLPRLVSTAIRASNSQNSADVERCPDVLRLKGLTASVKNLSHRSKVLPLESLDHNPNKIRIEAHVLFISWMTFSSGQSSSTCNKPNLKLKVVKACSELKRNEDLSHGFAVLKAKETVFGMKKNDERSRVETREMKOGDSKSPALMTIEKKQASPIQSGSKNPSRSGAGETGYIKLRLKPDCCSVSRPLMKKHDPVKAEPNPKTCIVYKRSPEMLFEFAKKSADYANAKGFLAAKEASICVDILMLKFSIISIAIERIRRMETLERITKDKDKICNAALALLHHNRQTRNQQ"

complement(1231..81848)

/note="Sequence from clone T21L14"

complement(join(<2300..3151,3250..3722,3900..>4203))

/gene="At2g32830"

complement(<2300..>4203)

/gene="At2g32830"

/note="T21L14.23"

complement(join(2300..3151,3250..3722,3900..4203))

/gene="At2g32830"

/codon\_start=1

/product="putative phosphate transporter"

/protein\_id="BAB91982.1"

/db\_xref="GI:2702279"

/translation="MAKKGKEVLNALDAKTOHYHTAIVIAGMGFETDAVDLFSISLVKLGRIVYHYDSSKRGTLPPNVAANGVAFGTLAGOLFPMGLGDKLRKRYGRTIAALVSGISGSGSPGHANGVMAVTFEPFPGIGDGPISLATISMYAKRTKAFIAAPAMGPGIILAGIYSLVSSFPDAFAKAPYENDPVSSTVPDQAYNRIVLMFGAIPALITTYWPKMKPETARKYALVALARNTKQASDKSKVLOVDLAEBAQSNSSNPPTFTGLFREFARRHGLHLTQTTTFLDLIAYSSNLFQDITYAIGIPAAETMAHAEVTVSGAQTALALCGTGVGYTFVAFIDILRFFIOLMGFTMTFMPALAIPYDHRHRENIIGLIMSLMEIPANGPNATITVPAEIPARLRSCYHISAAGKGAIVGAFGLFYAASSDSEKTDGPGICVRSILMLACVNLGIVFTLVPESKKSLIEISREDESGDPTVEMTANSGRKVPV"

5483..5600

/rpt\_family="(TAA)n"

complement(join(<7648..7741,7838..7917,8011..8118,8228..8305,8403..8442,8533..8585,8715..8747,8829)..>9383))

/gene="At2g32840"

complement(<7648..>9383)

/gene="At2g32840"

/note="T21L14.22: predicted by genefinder"

complement(join(7648..7741,7838..7917,8011..8118,8228..8305,8403..8442,8533..8585,8715..8747,8829)..9383))

/gene="At2g32840"

/note="hypothetical protein"

/codon\_start=1

/protein\_id="BAB91981.1"

/db\_xref="GI:2702278"

/translation="ASTDPDNPNPNPNSVSAVASTPIVTASESPVOTNPVITPSSQPOPQPASSYRALIAPLRHHPHONITVPLPRLRSNSYSTNSHOPHPDPSLIPFGSSGRFPTRPVARNNSVADVPYGSPPGCTPRGCVYGYHHQFVSNIDPMKQFRRAAHPQNOQSPOLGSHMKGVPHFLQPRVLVSENVVETAPSPSILDSNGHKHAKSRSDALVIVAKRRYRTTEGASLYSLCRSLRGAHGIKIRIDMCLPRLPVDYKRTESLPKDLVEPAICEDNEDESGYKHLSESDLLRHRIDRAKVBARLREERIKRIARYKARLALLPPEEOCRNE"

complement(join(9789..11142,11327..11576,11666..11716,11802..11967,12058..>12524))

/gene="At2g32850"

complement(9789..>12524)

/gene="At2g32850"

/note="T21L14.21"

complement(join(9895..11142,11327..11576,11666..11716,11802..11967,12058..12524))

/gene="At2g32850"

/codon\_start=1

/product="putative serine/threonine protein kinase"

/protein\_id="BAB91980.1"

/db\_xref="GI:2702277"

/translation="MMKTFPAQKPEAPLLEGRLFELGILKQVARNVIAEGGESSYILAQVNHASROYALKHMCIDDESLVEMKISVLSKSKHPNVVITYAGILDMGNKKBALMAFCGSLVDLENRGAYFEKQALITFDYDQNAVAFAMHOSPRIAHRDLKAENLILSSDGOMKLCDFGGSVSTNHKIFERAEEGIEEDENIRKYTPTPYRAPEMWDLFR

EMISEKDIWALCCLFRIICYEFINAFDEGSKMLTOIGNANRIPESPEYVIFITDLKEM  
LOASDPDRPDITOVKSOFLITVIFSSOLRSGKSMITTCOSANRKPSPAPRSPDRPPS  
SGSDSGGPGICAFATQAHKTATSVYSDPANNKMPEDPDSKSEKVPYRDSHOPKPS  
PYVGEARGIOIRNKDLETTISQKNTTPAANNMTRVSKDAFNSFVADPTTFTFDMKA  
PGAEALAEATIORLKDLETKQTKSEKAEITAFKFKLSAICORSOROEJODKOTLAKSA  
SPSPDRSSQONOPSQGMHSMSTPVSFLFVLVHLSDITKVEINYLILIMFLKOR  
DKHEGTWMELODRSNMSTGSDTNSMOFSDPAKVPMSASRKNNTINOQVTRRSK  
PAAAGTCGFEFPEFESFRAAATSAATASGTOBSGNSGNSSTQORGNKRENO  
KTRYPAGWAGF"

join(<14780, .15118,15202, .15271,15365, .15423,15533, .15608,  
15742, .15819,15934, .16021,16139, .16398,16541, .16655,  
16793, .17010,17342, .17479,17581, .17689,17849, .>18143)  
/gene=At2g32860  
<14780, .>18143  
/gene=At2g32860"  
/note="T21L14.20"  
join(14780, .15118,15202, .15271,15365, .15423,15533, .15608',  
15742, .15819,15934, .16021,16139, .16398,16541, .16655,  
16793, .17010,17342, .17479,17581, .17689,17849, .18143)  
/gene=At2g32860"  
/codon\_start=1  
/product="putative beta-glucosidase"  
/protein\_id="AB91979.1"  
/db\_xref="GI:2702276"  
/translation="MATATITLFLGLLALTSTILSNADARQPSDEDLGTTIGPHOT  
SFDELTATIGVPHATVDEDDIDMDKGTATGPOINLDDDLGTTIGPEFIHQDPAD  
FIFGTISAYQVEGARKSGRGLTISWDEFTIHMEPEKVOONGDEGVFTYRKDKID  
LMEELMTNGRFSISWTRILPGYTIKSGNEEGVKYNNLIELLANGIOCVTLFHK  
ESPLAEEMEGYCFINRIVEDREFRANCFQKREGEADKWAFTENESYVAGYKSK  
KAAGRSQWQAPKCPITGDSSEEPYIVAHNOILAHIAAVIDEPFNCKVEEGGKIGIVY  
SHMEPEKVSSEDKVARSRLSEYQGMFLRLPTVQGYAEMLDEVNITRLREFTEES  
EKLRKSLDEPYGLNYGAFESTPLAKVNSOQLNTEDLRVNMTYINNSLSPLQTTSM  
GIATYRGLKNIILKHKIDEYMDPEYIIMENGDEIDYGRKNTTEATINDYGRREFKSH  
LILMGKSRIMDKVRLKGYIWSLMDNFDEGKVFYGYVDYDNDMKRYTRSSGKW  
LSFLLDSKETLHKCFEGHREKGYAKLDFVELEIEMNSQLSYRSDFM"  
complement(18507, .19090)  
/note="molecular marker 96191"  
complement(join(19158, .19966,20113, .20413,20521, .20672,  
20766, .20895,20993, .21068))  
/gene=At2g32870"  
complement(19158, .21068)

gene

alignment\_scores:

Quality:	61.00	Length:	30
Ratio:	3.389	Gaps:	0
Percent Similarity:	60.000	Percent Identity:	36.667

alignment\_block:

US-09-331-631-5\_COPy\_1\_32 x AC003033 ..

Align seg 1/1 to: AC003033 from: 1 to: 84254

2 CysMetcGlnLeuGlnThrSerGlyGlnMetArgArgCysValSerGlnCyl 18  
||||:||||:||||: |||  
9237 TGTTCAGGTGACGCTGAGACGACCGCTATACACTGTCTGTTAGGTGTG 9286  
18 SasplysArpPheGluGluAspTrpSerLeuLys 31  
|.....|  
9287 TCACCGAGATTACATCGGTGACGATTGAGATCGATAC 9326